



-20-

SEQUENCE LISTING

#6

(1) GENERAL INFORMATION

- (i) APPLICANTS: Stefan Somlo and Toshio Mochizuki
- (ii) TITLE OF INVENTION: POLYCYSTIC KIDNEY DISEASE PKD2 GENE AND USES THEREOF
- (iii) NUMBER OF SEQUENCES: 12
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: AMSTER, ROTHSTEIN & EBENSTEIN
 - (B) STREET: 90 PARK AVENUE
 - (C) CITY: NEW YORK
 - (D) STATE: NEW YORK
 - (E) COUNTRY: U.S.A.
 - (F) ZIP: 10016
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: 3.5 INCH 1.44 Mb STORAGE DISKETTE
 - (B) COMPUTER: IBM PC COMPATIBLE
 - (C) OPERATING SYSTEM: MS-DOS
 - (D) SOFTWARE: ASCII
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER: 08/651,999
 - (B) FILING DATE: MAY 23, 1996
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: ELIZABETH A. BOGOSIAN
 - (B) REGISTRATION NUMBER: 39,911
 - (C) REFERENCE/DOCKET NUMBER: 96700/395
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: (212) 697-5995
 - (B) TELEFAX: (212) 286-0854 or 286-0082
 - (C) TELEX: TWX 710-581-4766

(2) INFORMATION FOR SEQ ID NO: 1

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 866 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE:
 - (A) DESCRIPTION: peptide
- (iii) HYPOTHETICAL: NO
- (v) FRAGMENT TYPE: internal fragment
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1

| | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|--|
| Met | Val | Asn | Ser | Ser | Arg | Val | Gln | Pro | Gln | | |
| 1 | | | | 5 | | | | | 10 | | |
| Gln | Pro | Gly | Asp | Ala | Lys | Arg | Pro | Pro | Ala | | |
| | | | | 15 | | | | | 20 | | |
| Pro | Arg | Ala | Pro | Asp | Pro | Gly | Arg | Leu | Met | | |
| | | | | 25 | | | | | 30 | | |
| Ala | Gly | Cys | Ala | Ala | Val | Gly | Ala | Ser | Leu | | |
| | | | | 35 | | | | | 40 | | |
| Ala | Ala | Pro | Gly | Gly | Leu | Cys | Glu | Gln | Arg | | |
| | | | | 45 | | | | | 50 | | |
| Gly | Leu | Glu | Ile | Glu | Met | Gln | Arg | Ile | Arg | | |
| | | | | 55 | | | | | 60 | | |
| Gln | Ala | Ala | Ala | Arg | Asp | Pro | Pro | Ala | Gly | | |
| | | | | 65 | | | | | 70 | | |
| Ala | Ala | Ala | Ser | Pro | Ser | Pro | Pro | Leu | Ser | | |
| | | | | 75 | | | | | 80 | | |
| Ser | Cys | Ser | Arg | Gln | Ala | Trp | Ser | Arg | Asp | | |
| | | | | 85 | | | | | 90 | | |

| | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|
| Asn | Pro | Gly | Glu | Glu | Glu | Ala | Glu | Glu | Glu | |
| | | | | 95 | | | | | 100 | |
| Glu | Glu | Glu | Val | Glu | Gly | Glu | Glu | Gly | Gly | |
| | | | | 105 | | | | | 110 | |
| Met | Val | Val | Glu | Met | Asp | Val | Glu | Trp | Arg | |
| | | | | 115 | | | | | 120 | |
| Pro | Gly | Ser | Arg | Arg | Ser | Ala | Ala | Ser | Ser | |
| | | | | 125 | | | | | 130 | |
| Ala | Val | Ser | Ser | Val | Gly | Ala | Arg | Ser | Arg | |
| | | | | 135 | | | | | 140 | |
| Gly | Leu | Gly | Gly | Tyr | His | Gly | Ala | Gly | His | |
| | | | | 145 | | | | | 150 | |
| Pro | Ser | Gly | Arg | Arg | Arg | Arg | Arg | Glu | Asp | |
| | | | | 155 | | | | | 160 | |
| Gln | Gly | Pro | Pro | Cys | Pro | Ser | Pro | Val | Gly | |
| | | | | 165 | | | | | 170 | |
| Gly | Gly | Asp | Pro | Leu | His | Arg | His | Leu | Pro | |
| | | | | 175 | | | | | 180 | |
| Leu | Glu | Gly | Gln | Pro | Pro | Arg | Val | Ala | Trp | |
| | | | | 185 | | | | | 190 | |
| Ala | Glu | Arg | Leu | Val | Arg | Gly | Leu | Arg | Gly | |
| | | | | 195 | | | | | 200 | |
| Leu | Trp | Gly | Thr | Arg | Leu | Met | Glu | Glu | Ser | |
| | | | | 205 | | | | | 210 | |
| Ser | Thr | Asn | Arg | Glu | Lys | Tyr | Leu | Lys | Ser | |
| | | | | 215 | | | | | 220 | |
| Val | Leu | Arg | Glu | Leu | Val | Thr | Tyr | Leu | Leu | |
| | | | | 225 | | | | | 230 | |

| | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|
| Phe | Leu | Ile | Val | Leu | Cys | Ile | Leu | Thr | Tyr | |
| | | | | 235 | | | | | 240 | |
| Gly | Thr | Glu | Ala | Asp | Asn | Arg | Ser | Phe | Ile | |
| | | | | 245 | | | | | 250 | |
| Phe | Tyr | Glu | Asn | Leu | Leu | Leu | Gly | Val | Pro | |
| | | | | 255 | | | | | 260 | |
| Arg | Ile | Arg | Gln | Leu | Arg | Val | Arg | Asn | Gly | |
| | | | | 265 | | | | | 270 | |
| Ser | Cys | Ser | Ile | Pro | Gln | Asp | Leu | Arg | Asp | |
| | | | | 275 | | | | | 280 | |
| Glu | Ile | Lys | Glu | Cys | Tyr | Asp | Val | Tyr | Glu | |
| | | | | 285 | | | | | 290 | |
| Thr | Ala | Ala | Gln | Val | Ala | Ser | Leu | Lys | Lys | |
| | | | | 295 | | | | | 300 | |
| Asn | Val | Trp | Leu | Asp | Arg | Gly | Thr | Arg | Ala | |
| | | | | 305 | | | | | 310 | |
| Thr | Phe | Ile | Asp | Phe | Ser | Val | Tyr | Asn | Ala | |
| | | | | 315 | | | | | 320 | |
| Asn | Ile | Asn | Leu | Phe | Cys | Val | Val | Arg | Leu | |
| | | | | 325 | | | | | 330 | |
| Leu | Val | Glu | Phe | Pro | Ala | Thr | Gly | Gly | Val | |
| | | | | 335 | | | | | 340 | |
| Ile | Pro | Ser | Trp | Gln | Phe | Gln | Pro | Leu | Lys | |
| | | | | 345 | | | | | 350 | |
| Leu | Ile | Arg | Tyr | Val | Thr | Thr | Phe | Asp | Phe | |
| | | | | 355 | | | | | 360 | |
| Phe | Leu | Ala | Ala | Cys | Glu | Ile | Ile | Phe | Cys | |
| | | | | 365 | | | | | 370 | |

| | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Phe | Phe | Ile | Phe | Tyr | Tyr | Val | Val | Glu | Glu |
| | | | | 375 | | | | | 380 |
| Ile | Leu | Glu | Xaa | Ile | Arg | Ile | His | Lys | Leu |
| | | | | 385 | | | | | 390 |
| His | Tyr | Phe | Arg | Xaa | Ser | Phe | Trp | Asn | Cys |
| | | | | 395 | | | | | 400 |
| Leu | Asp | Val | Val | Ile | Val | Val | Leu | Ser | Val |
| | | | | 405 | | | | | 410 |
| Val | Ala | Ile | Gly | Ile | Asn | Ile | Tyr | Arg | Thr |
| | | | | 415 | | | | | 420 |
| Ser | Asn | Val | Glu | Val | Xaa | Leu | Leu | Gln | Phe |
| | | | | 425 | | | | | 430 |
| Leu | Xaa | Glu | Asp | Gln | Asn | Thr | Phe | Pro | Asn |
| | | | | 435 | | | | | 440 |
| Phe | Glu | His | Leu | Ala | Tyr | Trp | Gln | Ile | Gln |
| | | | | 445 | | | | | 450 |
| Phe | Asn | Asn | Ile | Ala | Ala | Val | Thr | Val | Phe |
| | | | | 455 | | | | | 460 |
| Phe | Val | Trp | Ile | Lys | Leu | Phe | Lys | Phe | Ile |
| | | | | 465 | | | | | 470 |
| Asn | Phe | Asn | Arg | Thr | Met | Ser | Gln | Leu | Ser |
| | | | | 475 | | | | | 480 |
| Thr | Thr | Met | Ser | Arg | Cys | Ala | Lys | Asp | Leu |
| | | | | 485 | | | | | 490 |
| Phe | Gly | Phe | Ala | Ile | Met | Phe | Phe | Ile | Ile |
| | | | | 495 | | | | | 500 |
| Phe | Leu | Ala | Tyr | Ala | Gln | Leu | Ala | Tyr | Leu |
| | | | | 505 | | | | | 510 |

| | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Val | Phe | Gly | Thr | Gln | Val | Asp | Asp | Phe | Ser |
| | | | | 515 | | | | | 520 |
| Thr | Phe | Gln | Glu | Cys | Ile | Phe | Thr | Gln | Phe |
| | | | | 525 | | | | | 530 |
| Arg | Ile | Ile | Leu | Gly | Asp | Ile | Asn | Phe | Ala |
| | | | | 535 | | | | | 540 |
| Glu | Ile | Glu | Glu | Ala | Asn | Xaa | Arg | Val | Leu |
| | | | | 545 | | | | | 550 |
| Gly | Pro | Ile | Tyr | Phe | Thr | Thr | Phe | Val | Phe |
| | | | | 555 | | | | | 560 |
| Phe | Met | Phe | Phe | Ile | Leu | Leu | Asn | Met | Phe |
| | | | | 565 | | | | | 570 |
| Leu | Ala | Ile | Ile | Asn | Asp | Thr | Tyr | Ser | Glu |
| | | | | 575 | | | | | 580 |
| Val | Lys | Ser | Asp | Leu | Xaa | Xaa | Xaa | Ala | Gln |
| | | | | 585 | | | | | 590 |
| Gln | Lys | Ala | Glu | Met | Glu | Leu | Ser | Asp | Leu |
| | | | | 595 | | | | | 600 |
| Ile | Arg | Lys | Gly | Tyr | His | Lys | Ala | Leu | Val |
| | | | | 605 | | | | | 610 |
| Lys | Leu | Lys | Leu | Lys | Lys | Asn | Thr | Val | Asp |
| | | | | 615 | | | | | 620 |
| Asp | Ile | Ser | Glu | Ser | Leu | Arg | Gln | Gly | Gly |
| | | | | 625 | | | | | 630 |
| Gly | Lys | Leu | Asn | Phe | Asp | Glu | Leu | Arg | Gln |
| | | | | 635 | | | | | 640 |
| Asp | Leu | Lys | Gly | Lys | Gly | His | Thr | Asp | Ala |
| | | | | 645 | | | | | 650 |

| | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|
| Glu | Ile | Glu | Ala | Ile | Phe | Thr | Lys | Tyr | Asp | |
| | | | | 655 | | | | | 660 | |
| Gln | Asp | Gly | Asp | Gln | Glu | Leu | Thr | Glu | His | |
| | | | | 665 | | | | | 670 | |
| Glu | His | Gln | Gln | Met | Arg | Asp | Asp | Leu | Glu | |
| | | | | 675 | | | | | 680 | |
| Lys | Glu | Arg | Glu | Asp | Leu | Asp | Leu | Asp | His | |
| | | | | 685 | | | | | 690 | |
| Ser | Ser | Leu | Pro | Arg | Pro | Met | Ser | Ser | Arg | |
| | | | | 695 | | | | | 700 | |
| Ser | Phe | Pro | Arg | Ser | Leu | Asp | Asp | Ser | Glu | |
| | | | | 705 | | | | | 710 | |
| Glu | Asp | Asp | Asp | Glu | Asp | Ser | Gly | His | Ser | |
| | | | | 715 | | | | | 720 | |
| Ser | Arg | Arg | Arg | Gly | Ser | Ile | Ser | Ser | Gly | |
| | | | | 725 | | | | | 730 | |
| Val | Ser | Tyr | Glu | Glu | Phe | Gln | Val | Leu | Val | |
| | | | | 735 | | | | | 740 | |
| Arg | Arg | Val | Asp | Arg | Met | Glu | His | Ser | Ile | |
| | | | | 745 | | | | | 750 | |
| Gly | Ser | Ile | Val | Ser | Lys | Ile | Asp | Ala | Val | |
| | | | | 755 | | | | | 760 | |
| Ile | Val | Lys | Leu | Glu | Ile | Met | Glu | Arg | Ala | |
| | | | | 765 | | | | | 770 | |
| Lys | Leu | Lys | Arg | Arg | Glu | Val | Leu | Gly | Arg | |
| | | | | 775 | | | | | 780 | |
| Leu | Leu | Asp | Gly | Val | Ala | Glu | Asp | Glu | Arg | |
| | | | | 785 | | | | | 790 | |

| | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|
| Leu | Gly | Arg | Asp | Ser | Glu | Ile | His | Arg | Glu | |
| | | | | 795 | | | | | 800 | |
| Gln | Met | Glu | Arg | Leu | Val | Arg | Glu | Glu | Leu | |
| | | | | 805 | | | | | 810 | |
| Glu | Arg | Trp | Glu | Ser | Asp | Asp | Ala | Ala | Ser | |
| | | | | 815 | | | | | 820 | |
| Gln | Ile | Ser | His | Gly | Leu | Gly | Thr | Pro | Val | |
| | | | | 825 | | | | | 830 | |
| Gly | Leu | Asn | Gly | Gln | Pro | Arg | Pro | Arg | Ser | |
| | | | | 835 | | | | | 840 | |
| Ser | Arg | Pro | Ser | Ser | Ser | Gln | Ser | Xaa | Thr | |
| | | | | 845 | | | | | 850 | |
| Glu | Gly | Met | Glu | Gly | Ala | Gly | Gly | Asn | Gly | |
| | | | | 855 | | | | | 860 | |
| Ser | Ser | Asn | Val | His | Val | | | | | |
| | | | | 865 | | | | | | |

(3) INFORMATION FOR SEQ ID NO: 2

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 523 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE:
 - (A) DESCRIPTION: peptide
- (iii) HYPOTHETICAL: NO
- (v) FRAGMENT TYPE: internal fragment
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2

| | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|--|
| Phe | Leu | Ala | Lys | Glu | Glu | Ala | Arg | Lys | Val | | |
| 1 | | | | 5 | | | | | 10 | | |
| Lys | Arg | Leu | His | Gly | Met | Leu | Arg | Ser | Leu | | |
| | | | | 15 | | | | | 20 | | |
| Leu | Val | Tyr | Met | Leu | Phe | Leu | Leu | Val | Thr | | |
| | | | | 25 | | | | | 30 | | |
| Leu | Leu | Ala | Ser | Tyr | Gly | Asp | Ala | Ser | Cys | | |
| | | | | 35 | | | | | 40 | | |
| His | Gly | His | Ala | Tyr | Xaa | Arg | Leu | Gln | Ser | | |
| | | | | 45 | | | | | 50 | | |
| Xaa | Xaa | Xaa | Xaa | Xaa | Ala | Ile | Lys | Gln | Glu | | |
| | | | | 55 | | | | | 60 | | |
| Leu | His | Ser | Arg | Ala | Phe | Leu | Ala | Ile | Thr | | |
| | | | | 65 | | | | | 70 | | |
| Arg | Ser | Glu | Glu | Leu | Trp | Pro | Trp | Met | Ala | | |
| | | | | 75 | | | | | 80 | | |
| His | Val | Leu | Leu | Pro | Tyr | Val | His | Xaa | Xaa | | |
| | | | | 85 | | | | | 90 | | |
| Xaa | Xaa | Xaa | Gly | Asn | Gln | Ser | Ser | Pro | Glu | | |
| | | | | 95 | | | | | 100 | | |
| Xaa | Xaa | Xaa | Xaa | Xaa | Xaa | Xaa | Xaa | Xaa | Xaa | | |
| | | | | 105 | | | | | 110 | | |
| Xaa | Leu | Gly | Pro | Pro | Arg | Leu | Arg | Gln | Val | | |
| | | | | 115 | | | | | 120 | | |
| Arg | Leu | Gln | Glu | Ala | Leu | Tyr | Pro | Asp | Pro | | |
| | | | | 125 | | | | | 130 | | |
| Pro | Gly | Pro | Arg | Val | His | Thr | Cys | Ser | Ala | | |
| | | | | 135 | | | | | 140 | | |

| | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|
| Ala | Gly | Gly | Phe | Ser | Thr | Ser | Asp | Tyr | Asp | |
| | | | | 145 | | | | | 150 | |
| Val | Gly | Trp | Glu | Ser | Pro | His | Asn | Gly | Ser | |
| | | | | 155 | | | | | 160 | |
| Gly | Thr | Trp | Ala | Thr | Xaa | Xaa | Ser | Ala | Pro | |
| | | | | 165 | | | | | 170 | |
| Asp | Leu | Leu | Gly | Ala | Trp | Ser | Trp | Gly | Ser | |
| | | | | 175 | | | | | 180 | |
| Cys | Ala | Val | Tyr | Asp | Ser | Gly | Gly | Tyr | Val | |
| | | | | 185 | | | | | 190 | |
| Gln | Glu | Leu | Gly | Leu | Ser | Leu | Glu | Glu | Ser | |
| | | | | 195 | | | | | 200 | |
| Arg | Asp | Arg | Leu | Arg | Phe | Leu | Gln | Leu | His | |
| | | | | 205 | | | | | 210 | |
| Asn | Trp | Leu | Asp | Asn | Arg | Ser | Arg | Ala | Val | |
| | | | | 215 | | | | | 220 | |
| Phe | Leu | Glu | Leu | Thr | Arg | Tyr | Ser | Pro | Ala | |
| | | | | 225 | | | | | 230 | |
| Val | Gly | Leu | His | Ala | Ala | Val | Thr | Leu | Arg | |
| | | | | 235 | | | | | 240 | |
| Leu | Glu | Phe | Pro | Ala | Ala | Gly | Arg | Ala | Leu | |
| | | | | 245 | | | | | 250 | |
| Ala | Ala | Leu | Ser | Val | Arg | Pro | Phe | Ala | Leu | |
| | | | | 255 | | | | | 260 | |
| Arg | Arg | Leu | Ser | Ala | Gly | Leu | Ser | Leu | Pro | |
| | | | | 265 | | | | | 270 | |
| Leu | Leu | Thr | Ser | Val | Cys | Leu | Leu | Leu | Phe | |
| | | | | 275 | | | | | 280 | |

| | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|
| Ala | Val | His | Phe | Ala | Val | Ala | Glu | Ala | Arg | |
| | | | | 285 | | | | | 290 | |
| Thr | Trp | Xaa | His | Arg | Glu | Gly | Arg | Trp | Arg | |
| | | | | 295 | | | | | 300 | |
| Val | Leu | Arg | Leu | Gly | Ala | Trp | Ala | Arg | Trp | |
| | | | | 305 | | | | | 310 | |
| Leu | Leu | Val | Ala | Leu | Thr | Ala | Ala | Thr | Ala | |
| | | | | 315 | | | | | 320 | |
| Leu | Val | Arg | Leu | Ala | Gln | Leu | Gly | Ala | Ala | |
| | | | | 325 | | | | | 330 | |
| Asp | Arg | Gln | Xaa | Xaa | Trp | Thr | Arg | Phe | Val | |
| | | | | 335 | | | | | 340 | |
| Arg | Gly | Arg | Pro | Arg | Arg | Phe | Thr | Ser | Phe | |
| | | | | 345 | | | | | 350 | |
| Asp | Gln | Val | Ala | Gln | Leu | Ser | Ser | Ala | Ala | |
| | | | | 355 | | | | | 360 | |
| Arg | Gly | Leu | Ala | Ala | Ser | Leu | Leu | Phe | Leu | |
| | | | | 365 | | | | | 370 | |
| Leu | Leu | Val | Lys | Ala | Ala | Gln | Gln | Leu | Arg | |
| | | | | 375 | | | | | 380 | |
| Phe | Val | Arg | Gln | Trp | Ser | Val | Phe | Gly | Lys | |
| | | | | 385 | | | | | 390 | |
| Thr | Leu | Cys | Arg | Ala | Leu | Pro | Glu | Leu | Leu | |
| | | | | 395 | | | | | 400 | |
| Gly | Val | Thr | Leu | Gly | Leu | Val | Val | Leu | Gly | |
| | | | | 405 | | | | | 410 | |
| Val | Ala | Tyr | Ala | Gln | Leu | Ala | Ile | Leu | Leu | |
| | | | | 415 | | | | | 420 | |

| | | |
|-----------------------------------------|-----|-----|
| Val Ser Ser Cys Val Asp Ser Leu Trp Ser | 425 | 430 |
| Val Ala Gln Ala Leu Leu Xaa Xaa Xaa Xaa | 435 | 440 |
| Val Leu Cys Pro Gly Thr Gly Leu Ser Thr | 445 | 450 |
| Leu Cys Pro Ala Glu Ser Trp His Leu Ser | 455 | 460 |
| Pro Leu Leu Cys Val Gly Leu Trp Ala Leu | 465 | 470 |
| Arg Leu Trp Gly Ala Leu Arg Leu Gly Ala | 475 | 480 |
| Val Ile Leu Arg Trp Arg Tyr His Ala Leu | 485 | 490 |
| Arg Gly Glu Leu Tyr Arg Pro Ala Trp Glu | 495 | 500 |
| Pro Gln Asp Tyr Glu Met Val Glu Leu Phe | 505 | 510 |
| Xaa Xaa Xaa Xaa Xaa Xaa Xaa Leu Arg Arg | 515 | 520 |
| Leu Arg Leu | | |

(4) INFORMATION FOR SEQ ID NO: 3

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 399 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE:

- (A) DESCRIPTION: peptide
- (iii) HYPOTHETICAL: NO
- (v) FRAGMENT TYPE: internal fragment
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3

| | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|
| Glu | Asn | Arg | Lys | Met | Arg | Asp | Glu | Gln | Leu | |
| 1 | | | | 5 | | | | | 10 | |
| Phe | Ile | Thr | Ile | Arg | Asp | Met | Leu | Cys | Phe | |
| | | | | 15 | | | | | 20 | |
| Phe | Ala | Ser | Leu | Tyr | Ile | Met | Val | Met | Leu | |
| | | | | 25 | | | | | 30 | |
| Thr | Tyr | Tyr | Cys | Lys | Asp | Arg | His | Gly | Tyr | |
| | | | | 35 | | | | | 40 | |
| Trp | Tyr | Gln | Leu | Glu | Met | Ser | Thr | Ile | Leu | |
| | | | | 45 | | | | | 50 | |
| Asn | Ile | Asn | Gln | Lys | Asn | Tyr | Gly | Asp | Asn | |
| | | | | 55 | | | | | 60 | |
| Thr | Xaa | Phe | Met | Ser | Ile | Gln | His | Ala | Asp | |
| | | | | 65 | | | | | 70 | |
| Asp | Phe | Trp | Asp | Trp | Ala | Arg | Glu | Ser | Leu | |
| | | | | 75 | | | | | 80 | |
| Ala | Thr | Ala | Leu | Leu | Ala | Ser | Trp | Tyr | Asp | |
| | | | | 85 | | | | | 90 | |
| Gly | Asn | Pro | Ala | Tyr | Gly | Met | Arg | Ala | Tyr | |
| | | | | 95 | | | | | 100 | |
| Met | Asn | Asp | Lys | Val | Ser | Arg | Ser | Met | Gly | |
| | | | | 105 | | | | | 110 | |

| | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|--|
| Ile | Gly | Thr | Ile | Arg | Gln | Val | Arg | Thr | Lys | | |
| | | | | 115 | | | | | 120 | | |
| Lys | Ser | Glu | Ile | Ile | Thr | Leu | Phe | Asn | Lys | | |
| | | | | 125 | | | | | 130 | | |
| Leu | Asp | Ser | Glu | Arg | Trp | Ile | Asp | Asp | His | | |
| | | | | 135 | | | | | 140 | | |
| Thr | Arg | Ala | Val | Ile | Ile | Glu | Phe | Ser | Ala | | |
| | | | | 145 | | | | | 150 | | |
| Tyr | Asn | Ala | Gln | Ile | Asn | Tyr | Phe | Ser | Val | | |
| | | | | 155 | | | | | 160 | | |
| Val | Gln | Leu | Leu | Val | Glu | Ile | Pro | Lys | Ser | | |
| | | | | 165 | | | | | 170 | | |
| Gly | Ile | Tyr | Leu | Pro | Asn | Ser | Trp | Val | Glu | | |
| | | | | 175 | | | | | 180 | | |
| Ser | Val | Arg | Leu | Ile | Lys | Ser | Glu | Gly | Ser | | |
| | | | | 185 | | | | | 190 | | |
| Asp | Gly | Thr | Val | Val | Lys | Tyr | Tyr | Glu | Met | | |
| | | | | 195 | | | | | 200 | | |
| Leu | Tyr | Ile | Phe | Phe | Ser | Val | Leu | Ile | Phe | | |
| | | | | 205 | | | | | 210 | | |
| Val | Lys | Glu | Ile | Val | Trp | Asn | Phe | Met | Asp | | |
| | | | | 215 | | | | | 220 | | |
| Leu | Ile | Val | Gly | Ala | Leu | Ala | Val | Ala | Ser | | |
| | | | | 225 | | | | | 230 | | |
| Val | Leu | Ala | Tyr | Thr | Ile | Arg | Gln | Arg | Thr | | |
| | | | | 235 | | | | | 240 | | |
| Thr | Asn | Arg | Ala | Met | Glu | Asp | Phe | Asn | Ala | | |
| | | | | 245 | | | | | 250 | | |

| | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Asn | Asn | Gly | Asn | Ser | Tyr | Ile | Asn | Leu | Thr |
| | | | | 255 | | | | | 260 |
| Glu | Gln | Arg | Asn | Trp | Glu | Ile | Val | Phe | Ser |
| | | | | 265 | | | | | 270 |
| Tyr | Cys | Leu | Ala | Gly | Ala | Val | Phe | Phe | Thr |
| | | | | 275 | | | | | 280 |
| Ser | Cys | Lys | Met | Ile | Arg | Ile | Leu | Arg | Phe |
| | | | | 285 | | | | | 290 |
| Asn | Arg | Arg | Ile | Gly | Val | Leu | Ala | Ala | Thr |
| | | | | 295 | | | | | 300 |
| Leu | Asp | Asn | Ala | Leu | Gly | Ala | Ile | Val | Ser |
| | | | | 305 | | | | | 310 |
| Phe | Gly | Ile | Ala | Phe | Leu | Phe | Phe | Ser | Met |
| | | | | 315 | | | | | 320 |
| Thr | Phe | Asn | Ser | Val | Leu | Tyr | Ala | Val | Leu |
| | | | | 325 | | | | | 330 |
| Gly | Asn | Lys | Met | Gly | Gly | Tyr | Arg | Ser | Leu |
| | | | | 335 | | | | | 340 |
| Met | Ala | Thr | Phe | Gln | Thr | Ala | Leu | Ala | Gly |
| | | | | 345 | | | | | 350 |
| Met | Leu | Gly | Lys | Leu | Asp | Val | Thr | Ser | Ile |
| | | | | 355 | | | | | 360 |
| Gln | Pro | Xaa | Xaa | Xaa | Xaa | Xaa | Ile | Ser | Gln |
| | | | | 365 | | | | | 370 |
| Phe | Ala | Phe | Val | Val | Ile | Met | Leu | Tyr | Met |
| | | | | 375 | | | | | 380 |
| Ile | Glu | Phe | Glu | Glu | Ile | Arg | Asn | Asp | Ser |
| | | | | 385 | | | | | 390 |

Glu Lys Gln Thr Asn Asp Tyr Glu Ile
395

(5) INFORMATION FOR SEQ ID NO: 4

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 363 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE:
 - (A) DESCRIPTION: peptide
- (iii) HYPOTHETICAL: NO
- (v) FRAGMENT TYPE: internal fragment
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4

| | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Phe | Thr | Met | Val | Phe | Ser | Leu | Glu | Cys | Val |
| 1 | | | | 5 | | | | | 10 |
| Leu | Lys | Val | Ile | Ala | Phe | Gly | Phe | Leu | Asn |
| | | | | 15 | | | | | 20 |
| Tyr | Phe | Arg | Xaa | Asp | Thr | Trp | Asn | Ile | Phe |
| | | | | 25 | | | | | 30 |
| Asp | Phe | Ile | Thr | Val | Ile | Gly | Ser | Ile | Thr |
| | | | | 35 | | | | | 40 |
| Glu | Ile | Ile | Leu | Thr | Asp | Ser | Lys | Leu | Val |
| | | | | 45 | | | | | 50 |
| Asn | Thr | Ser | Gly | Xaa | Phe | Xaa | Xaa | Xaa | Xaa |
| | | | | 55 | | | | | 60 |
| Xaa | Xaa | Xaa | Xaa | Xaa | Xaa | Xaa | Xaa | Xaa | Xaa |
| | | | | 65 | | | | | 70 |

| | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Xaa | Asn | Met | Ser | Phe | Leu | Lys | Xaa | Xaa | Xaa | 75 | 80 |
| Xaa | Xaa | Xaa | Xaa | Xaa | Xaa | Xaa | Xaa | Leu | Phe | 85 | 90 |
| Arg | Ala | Ala | Arg | Leu | Ile | Lys | Leu | Leu | Arg | 95 | 100 |
| Gln | Gly | Tyr | Thr | Ile | Arg | Ile | Leu | Leu | Trp | 105 | 110 |
| Thr | Phe | Val | Gln | Ser | Phe | Lys | Ala | Leu | Pro | 115 | 120 |
| Tyr | Val | Cys | Leu | Leu | Ile | Ala | Met | Leu | Phe | 125 | 130 |
| Phe | Ile | Tyr | Ala | Ile | Ile | Gly | Met | Gln | Val | 135 | 140 |
| Phe | Gly | Asn | Asn | Phe | Arg | Ser | Phe | Phe | Gly | 145 | 150 |
| Ser | Leu | Met | Leu | Leu | Phe | Arg | Ser | Ala | Thr | 155 | 160 |
| Gly | Glu | Xaa | Ala | Trp | Gln | Glu | Ile | Glu | Arg | 165 | 170 |
| Cys | Gly | Xaa | Thr | Asp | Leu | Ala | Tyr | Val | Tyr | 175 | 180 |
| Phe | Val | Ser | Phe | Ile | Phe | Phe | Cys | Ser | Phe | 185 | 190 |
| Leu | Met | Leu | Asn | Leu | Phe | Val | Ala | Val | Ile | 195 | 200 |
| Met | Asp | Asn | Phe | Glu | Tyr | Leu | Thr | Arg | Asp | 205 | 210 |

| | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Xaa | Xaa | Xaa | Xaa | Xaa | Xaa | Xaa | Xaa | Xaa | Xaa |
| | | | | 215 | | | | | 220 |
| Xaa | Xaa | Xaa | Xaa | Xaa | Xaa | Xaa | Xaa | Xaa | Xaa |
| | | | | 225 | | | | | 230 |
| Xaa | Xaa | Xaa | Xaa | Xaa | Xaa | Xaa | Xaa | Xaa | Xaa |
| | | | | 235 | | | | | 240 |
| Xaa | Xaa | Xaa | Xaa | Xaa | Xaa | Xaa | Xaa | Xaa | Xaa |
| | | | | 245 | | | | | 250 |
| Xaa | Xaa | Xaa | Xaa | Xaa | Xaa | Xaa | Xaa | Xaa | Xaa |
| | | | | 255 | | | | | 260 |
| Xaa | Xaa | Xaa | Xaa | Xaa | Xaa | Xaa | Ile | Leu | Gly |
| | | | | 265 | | | | | 270 |
| Pro | His | His | Leu | Asp | Xaa | Glu | Phe | Val | Arg |
| | | | | 275 | | | | | 280 |
| Val | Trp | Ala | Glu | Tyr | Asp | Arg | Ala | Ala | Cys |
| | | | | 285 | | | | | 290 |
| Gly | Arg | Ile | His | Tyr | Thr | Glu | Met | Tyr | Glu |
| | | | | 295 | | | | | 300 |
| Met | Glu | Arg | Arg | Arg | Ser | Lys | Glu | Arg | Lys |
| | | | | 305 | | | | | 310 |
| His | Leu | Leu | Ser | Pro | Asp | Val | Ser | Arg | Cys |
| | | | | 315 | | | | | 320 |
| Asn | Ser | Glu | Glu | Arg | Gly | Thr | Gln | Ala | Asp |
| | | | | 325 | | | | | 330 |
| Trp | Glu | Ser | Pro | Glu | Arg | Arg | Gln | Ser | Arg |
| | | | | 335 | | | | | 340 |
| Ser | Pro | Ser | Glu | Gly | Arg | Ser | Gln | Thr | Pro |
| | | | | 345 | | | | | 350 |

Asn Arg Gln Gly Thr Gly Ser Leu Ser Glu
355 360

Ser Ser Ile

(4) INFORMATION FOR SEQ ID NO: 5

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 28 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE:
(A) DESCRIPTION: peptide

- (iii) HYPOTHETICAL: NO

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5

Glu Ile Ala Ile Phe Thr Lys Tyr Asp Gln
1 5 10

Asp Gly Asp Gln Glu Leu Thr Glu His Glu
15 20

His Gln Gln Met Arg Asp Asp Leu
25

(5) INFORMATION FOR SEQ ID NO: 6

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5057 nucleic acids
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE:
(A) DESCRIPTION: genomic DNA

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6

| | | | | | |
|------------|------------|-------------|------------|------------|------|
| GGCTCCTGAG | GCGCACAGCG | CCGAGCGCGG | CGCCGCGCAC | CCGCGCGCCG | 50 |
| GACGCCAGTG | ACCGCGATGG | TGAACTCCAG | TCGCGTGCAG | CCTCAGCAGC | 100 |
| CCGGGGACGC | CAAGCGGCCG | CCCGCGCCCC | GCGCGCCGGA | CCCGGGCCCG | 150 |
| CTGATGGCTG | GCTGCGCGGC | CGTGGGCGCC | AGCCTCGCCG | CCCCGGGCGG | 200 |
| CCTCTGCGAG | CAGCGGGGCC | TGGAGATCGA | GATGCAGCGC | ATCCGGCAGG | 250 |
| CGGCCGCGCG | GGACCCCCCG | GCCGGAGCCG | CGGCCTCCCC | TTCTCCTCCG | 300 |
| CTCTCGTCGT | GCTCCCGGCA | GCGGTGGAGC | CGCGATAACC | CCGGCTTCGA | 350 |
| GGCCGAGGAG | GAGGAGGAGG | AGGTGGAAGG | GGAAGAAGGC | GGAATGGTGG | 400 |
| TGGAGATGGA | CGTAGAGTGG | CGCCCGGGCA | GCCGGAGGTC | GGCCGCCTCC | 450 |
| TCGGCCGTGA | GCTCCGTGGG | CGCGCGGAGC | CGGGGGCTTG | GGGGCTACCA | 500 |
| CGGCGCGGGC | CACCCGAGCG | GGAGGCGGCG | CCGGCGAGAG | GACCAGGGCC | 550 |
| CGCCGTGCCC | CAGCCCAGTC | GGCGGCGGGG | ACCCGCTGCA | TCGCCACCTC | 600 |
| CCCCTGGAAG | GGCAGCCGCC | CCGAGTGGCC | TGGGCGGAGA | GGCTGGTTCC | 650 |
| CGGGCTGCGA | GGTCTCTGGG | GAACAAGACT | CATGGAGGAA | AGCAGCACTA | 700 |
| ACCGAGAGAA | ATACCTTAAA | AGTGTTTTAC | GGGAACTGGT | CACATACCTC | 750 |
| CTTTTTCTCA | TAGTCTTGTG | CATCTTGACC | TACGGCATGA | TGAGCTCCAA | 800 |
| TGTGTACTAC | TACACCCGGA | TGATGTCACA | GCTCTTCCTA | GACACCCCCG | 850 |
| TGTCCAAAAC | GGAGAAAAC | AACTTTAAAA | CTCTGTCTTC | CATGGAAGAC | 900 |
| TTCTGGAAGT | TCACAGAAGG | CTCCTTATTG | GATGGGCTGT | ACTGGAAGAT | 950 |
| GCAGCCCAGC | AACCAGACTG | AAGCTGACAA | CCGAAGTTTC | ATCTTCTATG | 1000 |
| AGAACCTGCT | GTTAGGGGTT | CCACGAATAC | GGCAACTCCG | AGTCAGAAAT | 1050 |
| GGATCCTGCT | CTATCCCCCA | GGACTTGAGA | GATGAAATTA | AAGAGTGCTA | 1100 |
| TGATGTCTAC | TCTGTCAGTA | GTGAAGATAG | GGCTCCCTTT | GGGCCCCGAA | 1150 |
| ATGGAACCGC | TTGGATCTAC | ACAAGTGAAA | AAGACTTGAA | TGGTAGTAGC | 1200 |
| CACTGGGGAA | TCATTGCAAC | TTATAGTGGA | GCTGGCTATT | ATCTGGATTT | 1250 |
| GTCAAGAACA | AGAGAGGAAA | CAGCTGCACA | AGTTGCTAGC | CTCAAGAAAA | 1300 |
| ATGTCTGGCT | GGACCGAGGA | ACCAGGGCAA | CTTTTATTGA | CTTCTCAGTG | 1350 |
| TACAACGCCA | ACATTAACCT | GTTCTGTGTG | GTCAGGTTAT | TGGTTGAATT | 1400 |
| CCCAGCAACA | GGTGGTGTGA | TTCCATCTTG | GCAATTCAG | CCTTTAAAGC | 1450 |
| TGATCCGATA | TGTCACAACT | TTTGATTTCT | TCCTGGCAGC | CTGTGAGATT | 1500 |
| ATCTTTTGTT | TCTTTATCTT | TTACTATGTG | GTGGAAGAGA | TATTGGAAAT | 1550 |
| TCGCATTAC | AAACTACACT | ATTTTCAGGAG | TTTCTGGAAT | TGTCTGGATG | 1600 |
| TTGTGATCGT | TGTGCTGTCA | GTGGTAGCTA | TAGGAATTAA | CATATACAGA | 1650 |
| ACATCAAATG | TGGAGGTGCT | ACTACAGTTT | CTGGAAGATC | AAAATACTTT | 1700 |
| CCCCAACTTT | GAGCATCTGG | CATATTGGCA | GATACAGTTC | AACAATATAG | 1750 |
| CTGCTGTCAC | AGTATTTTTT | GTCTGGATTA | AGCTCTTCAA | ATTCATCAAT | 1800 |
| TTTAACAGGA | CCATGAGCCA | GCTCTCGACA | ACCATGTCTC | GATGTGCCAA | 1850 |
| AGACCTGTTT | GGCTTTGCTA | TTATGTTCTT | CATTATTTTC | CTAGCGTATG | 1900 |

| | | | | | |
|-------------|-------------|------------|------------|------------|------|
| CTCAGTTGGC | ATACCTTGTC | TTTGGCACTC | AGGTCGATGA | CTTCAGTACT | 1950 |
| TTCCAAGAGT | GTATCTTCAC | TCAATTCCGT | ATCATTTTGG | GCGATATCAA | 2000 |
| CTTTGCAGAG | ATTGAGGAAG | CTAATCGAGT | TTTGGGACCA | ATTTATTTCA | 2050 |
| CTACATTTGT | GTTCTTTATG | TTCTTCATTC | TTTTGAATAT | GTTTTTGGCT | 2100 |
| ATCATCAATG | ATACTTACTC | TGAAGTGAAA | TCTGACTTGG | CACAGCAGAA | 2150 |
| AGCTGAAATG | GAACTCTCAG | ATCTTATCAG | AAAGGGCTAC | CATAAAGCTT | 2200 |
| TGGTCAAAC | AAAAC | AAAAATACCG | TGGATGACAT | TTCAGAGAGT | 2250 |
| CTGCGGCAAG | GAGGAGGCAA | GTTAAACTTT | GACGAACTTC | GACAAGATCT | 2300 |
| CAAAGGGAAG | GGCCATACTG | ATGCAGAGAT | TGAGGCAATA | TTCACAAAGT | 2350 |
| ACGACCAAGA | TGGAGACCAA | GAAGTACCG | AACATGAACA | TCAGCAGATG | 2400 |
| AGAGACGACT | TGGAGAAAGA | GAGGGAGGAC | CTGGATTTGG | ATCACAGTTC | 2450 |
| TTTACCACGT | CCCATGAGCA | GCCGAAGTTT | CCCTCGAAGC | CTGGATGACT | 2500 |
| CTGAGGAGGA | TGACGATGAA | GATAGCGGAC | ATAGCTCCAG | AAGGAGGGGA | 2550 |
| AGCATTTCTA | GTGGCGTTTC | TTACGAAGAG | TTTCAAGTCC | TGGTGAGACG | 2600 |
| AGTGGACCGG | ATGGAGCATT | CCATCGGCAG | CATAGTGTCC | AAGATTGACG | 2650 |
| CCGTGATCGT | GAAGCTAGAG | ATTATGGAGC | GAGCCAAACT | GAAGAGGAGG | 2700 |
| GAGGTGCTGG | GAAGGCTGTT | GGATGGGGTG | GCCGAGGATG | AAAGGCTGGG | 2750 |
| TCGTGACAGT | GAAATCCATA | GGGAACAGAT | GGAACGGCTA | GTACGTGAAG | 2800 |
| AGTTGGAACG | CTGGGAATCC | GATGATGCAG | CTTCCCAGAT | CAGTCATGGT | 2850 |
| TTAGGCACGC | CAGTGGGACT | AAATGGTCAA | CCTCGCCCCA | GAAGCTCCCG | 2900 |
| CCCATCTTCC | TCCCAATCTA | CAGAAGGCAT | GGAAGGTGCA | GGTGGAAATG | 2950 |
| GGAGTTCTAA | TGTCCACGTA | TGATATGTGT | GTTTCAGTAT | GTGTGTTTCT | 3000 |
| AATAAGTGAG | GAAGTGGCTG | TCCTGAATTG | CTGTAACAAG | CACACTATTT | 3050 |
| ATATGCCCTG | ACCACCATAG | GATGCTAGTC | TTTGTGACCG | ATTGCTAATC | 3100 |
| TTCTGCACTT | TAATTTATTT | TATATAAACT | TTACCCATGG | TTCAAAGATT | 3150 |
| TTTTTTTTCTT | TTTCTCATAT | AAGAAATCTA | GGTGTAATA | TTGAGTACAG | 3200 |
| AAAAAAAATC | TTCATGATGT | GTATTGAGCG | GTACGCCCAG | TTGCCACCAT | 3250 |
| GACTGAGTCT | TCTCAGTTGA | CAATGAAGTA | GCCTTTTAAA | GCTAGAAAAC | 3300 |
| TGTCAAAGGG | CTTCTGAGTT | TCATTTCCAG | TCACAAAAAT | CAGTATTGTT | 3350 |
| ATTTTTTTTCC | AAGAGTGTGA | AGGAAAATGG | GGCAATTCCT | TTCCACTCTG | 3400 |
| GCATAGTTCA | TGAGCTTAAT | ACATAGCTTT | CTTTTAAGAA | AGGAGCCTTT | 3450 |
| TTTTTCAACT | AGCTTCCTGG | GGTAAACTTT | TCTAAAAGAT | AAAATGGGAA | 3500 |
| GGAAC | ACTATGATAG | AATCTGTGTG | AATGGTTAAG | ATGAATGTTA | 3550 |
| AATACTATGC | TTTTTTGTAA | GTTGATCGTA | TCTGATGTCT | GTGGGACTAA | 3600 |
| CTGTATCACT | TAATTTTTTAC | CTTATTTTGG | CTCTAATTTG | AATAAGCTGA | 3650 |
| GTAAAACCAC | CAAAGATCAG | TTATAGGATA | AAATGGCATC | TCTAACCATA | 3700 |
| ACACAGGAGA | ATTGGAAGGA | GCCCTAAGTT | GTCCTCAGT | TTAATTTCTT | 3750 |
| TTAATGGTTA | GTTTAGCCTA | AAGATTTATC | TGCATATTCT | TTTTCCCATG | 3800 |
| TGGCTCTACT | CATTTGCAAC | TGAATTTAAT | GTTATAACTC | ATCTAGTGAG | 3850 |
| ACCAACTTAC | TAAATTTTTA | GTATGCACTG | AAAGTTTTTA | TCCAACAATT | 3900 |
| ATGTTCAATTT | TAAGCAAAAT | TTTAAGAAAG | TTTTGAAATT | CATAAAGCAT | 3950 |
| TTGTTTTTAA | ACTATTTTAA | GAATATAGTA | CTCGGTCAGG | TATGNNNCAC | 4000 |

| | | | | | |
|------------|------------|------------|------------|------------|------|
| GCCTGTAATC | CCAGCACTTT | GGGAGGCCGA | AACAGGCCGA | TCACCTGAGC | 4050 |
| CCAGGAGTTC | AAGACCAACA | TGGGCAATGT | GGCGAAACTC | CATCTCTACA | 4100 |
| AAAAATGCAA | AAATAAAAAA | TATAGTACTC | AAGTATTCTT | GATCCTGTGT | 4150 |
| TTCAAAACTA | GAATTTGTAA | TGCAAATGGA | GCTCAGTCTA | ATAAAAAAGA | 4200 |
| GGTTTTGGTA | TTAAAAGTTC | ATACATTAGA | CAGTATCAGC | CAAAATTTGA | 4250 |
| GTTAGCAACA | CTGTTTTCTT | TACGAGAGGG | TCTCACCCAA | ATTTATGGGG | 4300 |
| AGAAATCTAT | TTCTCAAAAA | AAAAAAATCT | TCTTTTACAG | AAATGTTGAG | 4350 |
| TAAGGTGACA | TTTTGAGCGC | TAATAAGCAA | AAGAGCATGC | AGTGCTGTTG | 4400 |
| AATAACCCTC | ACTTGGAGAA | CCAAGAGAAT | CCTGTCGTTT | AATGCTATAT | 4450 |
| TTTAATTTCA | CAAGTTGTTT | ATTTAACTGG | TAGAATGTCA | GTCCAATCTC | 4500 |
| CAATGAGAAC | ATGAGCAAAT | AGACCTTTCC | AGGTTGAAAG | TGAAACATAC | 4550 |
| TGGGTTTCTG | TAAGTTTTTC | CTCATGGCTT | CATCTCTATC | TTTACTTTCT | 4600 |
| CTTGAATATG | CTACACAAAG | TTCTTTATTA | CTACATACTA | AAGTTTGCAT | 4650 |
| TCCAGGGATA | TTGACTGTAC | ATATTTATGT | ATATGTACCA | TGTTGTTACA | 4700 |
| TGTAAACAAA | CTTCAATTTG | AAGTGCAGCT | ATTATGTGGT | ATCCATGTGT | 4750 |
| ATCGACCATG | TGCCATATAT | CAATTATGGT | CACTAGAAAG | TCTCTTTATG | 4800 |
| ATACTTTTTA | TTGTACTGTT | TTTCATTTCA | CTTGCAAAAT | TTTGCAGAAT | 4850 |
| TCCTCCTTTC | TACCCATAAA | TTACATATAA | TTTTTCTTCT | TTAGTCATGG | 4900 |
| AGAACNCCCC | CCCATCATCT | CANCCCTATT | ANCTTTCCCA | TGTGTACTGG | 4950 |
| TATTATTAAA | AAGACATTTA | CATACGCAAG | TTTTTCACTG | ACAANCAAGA | 5000 |
| ATGTTATTAA | TGTGTAATAC | TGAGCACNTT | TACTTCTTAA | TAAAAACTTG | 5050 |
| ATATANT | | | | | 5057 |

(5) INFORMATION FOR SEQ ID NO: 7

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 968 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE:

- (A) DESCRIPTION: peptide

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Val | Asn | Ser | Ser | Arg | Val | Gln | Pro | Gln | Gln | Pro | Gly | Asp | Ala | Lys |
| 1 | | | | 5 | | | | 10 | | | | | | 15 | |

| | | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|--|--|--|
| Arg | Pro | Pro | Ala | Pro | Arg | Ala | Pro | Asp | Pro | Gly | Arg | Leu | Met | Ala | Gly | | | | |
| | | | 20 | | | | | 25 | | | | | 30 | | | | | | |
| Cys | Ala | Ala | Val | Gly | Ala | Ser | Leu | Ala | Ala | Pro | Gly | Gly | Leu | Cys | Glu | | | | |
| | | 35 | | | | | 40 | | | | | 45 | | | | | | | |
| Gln | Arg | Gly | Leu | Glu | Ile | Glu | Met | Gln | Arg | Ile | Arg | Gln | Ala | Ala | Ala | | | | |
| | 50 | | | | | | 55 | | | | | 60 | | | | | | | |
| Arg | Asp | Pro | Pro | Ala | Gly | Ala | Ala | Ala | Ser | Pro | Ser | Pro | Pro | Leu | Ser | | | | |
| 65 | | | | | 70 | | | | | 75 | | | | | 80 | | | | |
| Ser | Cys | Ser | Arg | Gln | Ala | Trp | Ser | Arg | Asp | Asn | Pro | Gly | Phe | Glu | Ala | | | | |
| | | | | 85 | | | | | 90 | | | | | 95 | | | | | |
| Glu | Glu | Glu | Glu | Glu | Glu | Val | Glu | Gly | Glu | Glu | Gly | Gly | Met | Val | Val | | | | |
| | | | | 100 | | | | 105 | | | | | 110 | | | | | | |
| Glu | Met | Asp | Val | Glu | Trp | Arg | Pro | Gly | Ser | Arg | Arg | Ser | Ala | Ala | Ser | | | | |
| | | 115 | | | | | 120 | | | | | 125 | | | | | | | |
| Ser | Ala | Val | Ser | Ser | Val | Gly | Ala | Arg | Ser | Arg | Gly | Leu | Gly | Gly | Tyr | | | | |
| | 130 | | | | | 135 | | | | | 140 | | | | | | | | |
| His | Gly | Ala | Gly | His | Pro | Ser | Gly | Arg | Arg | Arg | Arg | Arg | Glu | Asp | Gln | | | | |
| 145 | | | | | 150 | | | | | 155 | | | | | 160 | | | | |
| Gly | Pro | Pro | Cys | Pro | Ser | Pro | Val | Gly | Gly | Gly | Asp | Pro | Leu | His | Arg | | | | |
| | | | | 165 | | | | | 170 | | | | | 175 | | | | | |
| His | Leu | Pro | Leu | Glu | Gly | Gln | Pro | Pro | Arg | Val | Ala | Trp | Ala | Glu | Arg | | | | |
| | | | 180 | | | | | 185 | | | | | 190 | | | | | | |
| Leu | Val | Arg | Gly | Leu | Arg | Gly | Leu | Trp | Gly | Thr | Arg | Leu | Met | Glu | Glu | | | | |
| | | 195 | | | | | 200 | | | | | 205 | | | | | | | |
| Ser | Ser | Thr | Asn | Arg | Glu | Lys | Tyr | Leu | Lys | Ser | Val | Leu | Arg | Glu | Leu | | | | |
| | 210 | | | | | 215 | | | | | 220 | | | | | | | | |
| Val | Thr | Tyr | Leu | Leu | Phe | Leu | Ile | Val | Leu | Cys | Ile | Leu | Thr | Tyr | Gly | | | | |
| 225 | | | | | 230 | | | | | 235 | | | | | 240 | | | | |

| | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|--|--|
| Met | Met | Ser | Ser | Asn | Val | Tyr | Tyr | Tyr | Thr | Arg | Met | Met | Ser | Gln | Leu | | | |
| | | | | 245 | | | | | 250 | | | | | 255 | | | | |
| Phe | Leu | Asp | Thr | Pro | Val | Ser | Lys | Thr | Glu | Lys | Thr | Asn | Phe | Lys | Thr | | | |
| | | | 260 | | | | | 265 | | | | | 270 | | | | | |
| Leu | Ser | Ser | Met | Glu | Asp | Phe | Trp | Lys | Phe | Thr | Glu | Gly | Ser | Leu | Leu | | | |
| | | 275 | | | | | 280 | | | | | 285 | | | | | | |
| Asp | Gly | Leu | Tyr | Trp | Lys | Met | Gln | Pro | Ser | Asn | Gln | Thr | Glu | Ala | Asp | | | |
| | 290 | | | | | 295 | | | | | 300 | | | | | | | |
| Asn | Arg | Ser | Phe | Ile | Phe | Tyr | Glu | Asn | Leu | Leu | Leu | Gly | Val | Pro | Arg | | | |
| 305 | | | | 310 | | | | | 315 | | | | | 320 | | | | |
| Ile | Arg | Gln | Leu | Arg | Val | Arg | Asn | Gly | Ser | Cys | Ser | Ile | Pro | Gln | Asp | | | |
| | | | 325 | | | | | 330 | | | | | | 335 | | | | |
| Leu | Arg | Asp | Glu | Ile | Lys | Glu | Cys | Tyr | Asp | Val | Tyr | Ser | Val | Ser | Ser | | | |
| | | 340 | | | | | 345 | | | | | | 350 | | | | | |
| Glu | Asp | Arg | Ala | Pro | Phe | Gly | Pro | Arg | Asn | Gly | Thr | Ala | Trp | Ile | Tyr | | | |
| | 355 | | | | | 360 | | | | | 365 | | | | | | | |
| Thr | Ser | Glu | Lys | Asp | Leu | Asn | Gly | Ser | Ser | His | Trp | Gly | Ile | Ile | Ala | | | |
| | 370 | | | | 375 | | | | | 380 | | | | | | | | |
| Thr | Tyr | Ser | Gly | Ala | Gly | Tyr | Tyr | Leu | Asp | Leu | Ser | Arg | Thr | Arg | Glu | | | |
| 385 | | | | 390 | | | | 395 | | | | | | 400 | | | | |
| Glu | Thr | Ala | Ala | Gln | Val | Ala | Ser | Leu | Lys | Lys | Asn | Val | Trp | Leu | Asp | | | |
| | | | 405 | | | | | 410 | | | | | | 415 | | | | |
| Arg | Gly | Thr | Arg | Ala | Thr | Phe | Ile | Asp | Phe | Ser | Val | Tyr | Asn | Ala | Asn | | | |
| | | | 420 | | | | | 425 | | | | | 430 | | | | | |
| Ile | Asn | Leu | Phe | Cys | Val | Val | Arg | Leu | Leu | Val | Glu | Phe | Pro | Ala | Thr | | | |
| | 435 | | | | | 440 | | | | | 445 | | | | | | | |
| Gly | Gly | Val | Ile | Pro | Ser | Trp | Gln | Phe | Gln | Pro | Leu | Lys | Leu | Ile | Arg | | | |
| | 450 | | | | | 455 | | | | | 460 | | | | | | | |

| | | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Tyr | Val | Thr | Thr | Phe | Asp | Phe | Phe | Leu | Ala | Ala | Cys | Glu | Ile | Ile | Phe | 465 | 470 | 475 | 480 |
| Cys | Phe | Phe | Ile | Phe | Tyr | Tyr | Val | Val | Glu | Glu | Ile | Leu | Glu | Ile | Arg | 485 | 490 | 495 | |
| Ile | His | Lys | Leu | His | Tyr | Phe | Arg | Ser | Phe | Trp | Asn | Cys | Leu | Asp | Val | 500 | 505 | 510 | |
| Val | Ile | Val | Val | Leu | Ser | Val | Val | Ala | Ile | Gly | Ile | Asn | Ile | Tyr | Arg | 515 | 520 | 525 | |
| Thr | Ser | Asn | Val | Glu | Val | Leu | Leu | Gln | Phe | Leu | Glu | Asp | Gln | Asn | Thr | 530 | 535 | 540 | |
| Phe | Pro | Asn | Phe | Glu | His | Leu | Ala | Tyr | Trp | Gln | Ile | Gln | Phe | Asn | Asn | 545 | 550 | 555 | 560 |
| Ile | Ala | Ala | Val | Thr | Val | Phe | Phe | Val | Trp | Ile | Lys | Leu | Phe | Lys | Phe | 565 | 570 | 575 | |
| Ile | Asn | Phe | Asn | Arg | Thr | Met | Ser | Gln | Leu | Ser | Thr | Thr | Met | Ser | Arg | 580 | 585 | 590 | |
| Cys | Ala | Lys | Asp | Leu | Phe | Gly | Phe | Ala | Ile | Met | Phe | Phe | Ile | Ile | Phe | 595 | 600 | 605 | |
| Leu | Ala | Tyr | Ala | Gln | Leu | Ala | Tyr | Leu | Val | Phe | Gly | Thr | Gln | Val | Asp | 610 | 615 | 620 | |
| Asp | Phe | Ser | Thr | Phe | Gln | Glu | Cys | Ile | Phe | Thr | Gln | Phe | Arg | Ile | Ile | 625 | 630 | 635 | 640 |
| Leu | Gly | Asp | Ile | Asn | Phe | Ala | Glu | Ile | Glu | Glu | Ala | Asn | Arg | Val | Leu | 645 | 650 | 655 | |
| Gly | Pro | Ile | Tyr | Phe | Thr | Thr | Phe | Val | Phe | Phe | Met | Phe | Phe | Ile | Leu | 660 | 665 | 670 | |
| Leu | Asn | Met | Phe | Leu | Ala | Ile | Ile | Asn | Asp | Thr | Tyr | Ser | Glu | Val | Lys | 675 | 680 | 685 | |

| | | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Ser | Asp | Leu | Ala | Gln | Gln | Lys | Ala | Glu | Met | Glu | Leu | Ser | Asp | Leu | Ile | 690 | 695 | 700 | |
| Arg | Lys | Gly | Tyr | His | Lys | Ala | Leu | Val | Lys | Leu | Lys | Leu | Lys | Lys | Asn | 705 | 710 | 715 | 720 |
| Thr | Val | Asp | Asp | Ile | Ser | Glu | Ser | Leu | Arg | Gln | Gly | Gly | Gly | Lys | Leu | 725 | 730 | 735 | |
| Asn | Phe | Asp | Glu | Leu | Arg | Gln | Asp | Leu | Lys | Gly | Lys | Gly | His | Thr | Asp | 740 | 745 | 750 | |
| Ala | Glu | Ile | Glu | Ala | Ile | Phe | Thr | Lys | Tyr | Asp | Gln | Asp | Gly | Asp | Gln | 755 | 760 | 765 | |
| Glu | Leu | Thr | Glu | His | Glu | His | Gln | Gln | Met | Arg | Asp | Asp | Leu | Glu | Lys | 770 | 775 | 780 | |
| Glu | Arg | Glu | Asp | Leu | Asp | Leu | Asp | His | Ser | Ser | Leu | Pro | Arg | Pro | Met | 785 | 790 | 795 | 800 |
| Ser | Ser | Arg | Ser | Phe | Pro | Arg | Ser | Leu | Asp | Asp | Ser | Glu | Glu | Asp | Asp | 805 | 810 | 815 | |
| Asp | Glu | Asp | Ser | Gly | His | Ser | Ser | Arg | Arg | Arg | Gly | Ser | Ile | Ser | Ser | 820 | 825 | 830 | |
| Gly | Val | Ser | Tyr | Glu | Glu | Phe | Gln | Val | Leu | Val | Arg | Arg | Val | Asp | Arg | 835 | 840 | 845 | |
| Met | Glu | His | Ser | Ile | Gly | Ser | Ile | Val | Ser | Lys | Ile | Asp | Ala | Val | Ile | 850 | 855 | 860 | |
| Val | Lys | Leu | Glu | Ile | Met | Glu | Arg | Ala | Lys | Leu | Lys | Arg | Arg | Glu | Val | 865 | 870 | 875 | 880 |
| Leu | Gly | Arg | Leu | Leu | Asp | Gly | Val | Ala | Glu | Asp | Glu | Arg | Leu | Gly | Arg | 885 | 890 | 895 | |
| Asp | Ser | Glu | Ile | His | Arg | Glu | Gln | Met | Glu | Arg | Leu | Val | Arg | Glu | Glu | 900 | 905 | 910 | |

Leu Glu Arg Trp Glu Ser Asp Asp Ala Ala Ser Gln Ile Ser His Gly
915 920 925

Leu Gly Thr Pro Val Gly Leu Asn Gly Gln Pro Arg Pro Arg Ser Ser
930 935 940

Arg Pro Ser Ser Ser Gln Ser Thr Glu Gly Met Glu Gly Ala Gly Gly
945 950 955 960

Asn Gly Ser Ser Asn Val His Val
965

(7) INFORMATION FOR SEQ ID NO: 8

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 nucleic acids
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE:

- (A) DESCRIPTION: genomic DNA

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8

GGGCTACCAT AAAGCTTTG

19

(8) INFORMATION FOR SEQ ID NO: 9

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 nucleic acids
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE:

- (A) DESCRIPTION: genomic DNA

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9

GTTCATGTTC GATCAGTTCT 20

(9) INFORMATION FOR SEQ ID NO: 10

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 23 nucleic acids

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE:

(A) DESCRIPTION: genomic DNA

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10

GGGCTAGAAA TACTCTTATC ACC 23

(10) INFORMATION FOR SEQ ID NO: 11

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 nucleic acids

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE:

(A) DESCRIPTION: genomic DNA

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11

GCCTCAAGTG TTCCACTGAT 20

(11) INFORMATION FOR SEQ ID NO: 12

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 22 nucleic acids
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE:
 - (A) DESCRIPTION: genomic DNA
- (iii) HYPOTHETICAL: NO
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12

AGGTTTTTCT GGGTAACCCT AG 22